

A medical multimodal large language model for future pandemics

Year: 2023 | Citations: 151 | Authors: Fenglin Liu, Tingting Zhu, Xian Wu

Abstract

Deep neural networks have been integrated into the whole clinical decision procedure which can improve the efficiency of diagnosis and alleviate the heavy workload of physicians. Since most neural networks are supervised, their performance heavily depends on the volume and quality of available labels. However, few such labels exist for rare diseases (e.g., new pandemics). Here we report a medical multimodal large language model (Med-MLLM) for radiograph representation learning, which can learn broad medical knowledge (e.g., image understanding, text semantics, and clinical phenotypes) from unlabelled data. As a result, when encountering a rare disease, our Med-MLLM can be rapidly deployed and easily adapted to them with limited labels. Furthermore, our model supports medical data across visual modality (e.g., chest X-ray and CT) and textual modality (e.g., medical report and free-text clinical note); therefore, it can be used for clinical tasks that involve both visual and textual data. We demonstrate the effectiveness of our Med-MLLM by showing how it would perform using the COVID-19 pandemic “in replay”. In the retrospective setting, we test the model on the early COVID-19 datasets; and in the prospective setting, we test the model on the new variant COVID-19-Omicron. The experiments are conducted on 1) three kinds of input data; 2) three kinds of downstream tasks, including disease reporting, diagnosis, and prognosis; 3) five COVID-19 datasets; and 4) three different languages, including English, Chinese, and Spanish. All experiments show that our model can make accurate and robust COVID-19 decision-support with little labelled data.